CZECH POLAR REPORTS 5 (2): 143-159, 2015

Exploitation of databases in polar research - Data evaluation and outputs

Jana Kvíderová[1,2*], Josef Elster[1,2], Ivan Iliev[3]

[1]*Centre of Polar Ecology, Faculty of Science, University of South Bohemia, Na Zlaté stoce 3, 370 05 České Budějovice, Czech Republic*
[2]*Institute of Botany AS CR, Dukelská 135, Třeboň, Czech Republic*
[3]*Institute of Plant Physiology and Genetics, Bulgarian Academy of Sciences, Acad. G. Bonchev Street, Bldg. 21, 1113 Sofia, Bulgaria*

Abstract

The increasing number of observations and floristic sample analyses provided by the Centre for Polar Ecology, Faculty of Science, University of South Bohemia in České Budějovice, Czech Republic (CPE), led to development of the sample database (SampleDTB). At present, the Sample DTB contains records on total of 318 samples from 135 sites. Total of 254 taxa at level of genera or species were observed. For database functionality tests, two datasets were selected. The first one consisted of samples collected by ALGO groups in frame of Polar Ecology course organized by the CPE in 2011-2014 (ALGO dataset). The second one consisted of samples collected in Bulgaria in 2013 (BG dataset). The ALGO dataset contains records on total of 188 samples from 94 sites. Total of 216 taxa (genera or species) were observed. The majority of habitats sampled were streams and the most frequently sampled communities were layer-forming communities like periphyton, epilithon and biofilms. The most dominant classes observed in ALGO dataset were diatoms and cyanobacteria. The unconstrained canonical analysisrevealed that the environment type significantly affected the taxonomical composition observed. In the BG dataset, 16 samples from 4 sites were recorded. Total of 40 taxa (genera or species) were observed. Majority of habitats sampled included streams and shallow pools and the communities sampled were restricted to periphyton, epiphyton and epilithon. Similarly to ALGO dataset, the most

---

Received October 30, 2015, *accepted* December 5, 2015.

[*] Corresponding author: Jana Kvíderová <jana.kviderova@objektivem.net>

Acknowledgements: The study was realized within the project *Creating of Working Team and Pedagogical Conditions for Teaching and Education in the Field of Polar Ecology and Life in Extreme Environment*, reg. No. CZ.1.07/2.2.00/28.0190 co-financed by the European Social Fund and the state budget of the Czech Republic. The research was also supported by projects LM2015078 - *CzechPolar2 - Czech Polar Research Infrastructure* (MŠMT), as priority topic of scientific cooperation between ASCR and BAS 2014 – 2016 No. 23 *Study of biotechnological potential of extremophilic and extremotolerant algae and cyanobacteria*, and a long-term research development project no. RVO 67985939 (IB). We would like to thank all people who helped during sampling, microscopical analyses and taxa identifications.

- Employees and students of the CPE (in alphabetical order): Josef Elster, Jiří Komárek, Jana Kvíderová, Otakar Strunecký, Katya Pushkareva, Daria Tashyreva, Lukáš Veselý, Jakub Žárský;
- Groups ALGO2011, ALGO2012, ALGO2013 and ALGO2014 (see reports on Polar Ecology course at polar.prj.jcu.cz for group members lists);
- Ivan Iliev, Institute of Plant Physiology and Genetics, Bulgarian Academy of Sciences, Sofia, BG.

dominant classes in BG dataset were Cyanophyceae and diatoms. Due to low number of samples, no multivariate analysis was performed in the BG dataset. These exports and further analyses proved functionality of the SampleDTB database.

Key words: database, data utilisation, species diversity, ecology

DOI: 10.5817/CPR2015-2-13

Introduction

At present, the importance of databases or data warehouses increases, since there is a pressure for data sharing, either from grant agencies or leading journals (Khalsa et Yarmey 2014). For project-specific requirements, a relational database are more suitable, however, sharing data in this type of databases is constrained in several ways like software used (Jones et al. 2006). Despite of their limitations, the relational databases containing ecological or ecophysiological data represent a unique source to assess diversity and ecosystem dynamics, even at a long-term scale. Moreover, they allow generation of reports and data exports for further analyses, like diversity indices calculations or multivariate analysis. Finally, they provide data for future resampling of a given site or for searching specific taxa, allowing thus long-term comparisons.

In polar sciences, the long-term repeated measurements of selected sites stored in databases could provide a valuable tool to reveal possible changes or trends in spatial/temporal changes in polar habitats caused by global change. These data may be also used in models of possible effects of climate change in the polar ecosystem. For instance, the measurements of algal growth potential may be used as a proxy of available nutrient concentration in polar streams, and therefore primary production of photosynthetic microorganisms, which could be influenced by increased weathering rate due to temperature rise (Kvíderová et Elster 2013).

For polar microbiology research, the database should store data on samples, conditions during sampling, taxa observed as well as results of laboratory analyses performed later. The database should contain links to supplementary data like site and community photographs and microphotographs of observed taxa. Such database, SampleDTB, is used the by CPE. The required outputs include number of samples collected, number and list of sites visited, number and list of taxa observed, number of habitats and communities sampled. For multivariate analysis, the database should be able to export environmental and species composition data in each sample formatted for easy export to table processor, like Excell® (Kvíderová 2014).

For database functionality tests, data collected in the frame of Polar Ecology courses organized by CPE in 2011 – 2014 and some extreme localities in Bulgaria in 2013 were used. The sample data were entered manually, rewritten from Sampling Protocol notebooks. The database tests were focused on data export for required outputs (numbers and lists of sites visited, samples collected, taxa observed, habitat and communities sampled) and their exploitation in multivariate analysis.

Material and Methods

Database structure and data sources

The SampleDTB is a project-specific relational database (refer to Kvíderová 2014 for its structure and data management). The data included in the SampleDTB were collected by a large group of people (*see* Acknowledgements).

The data for evaluation were divided into two datasets. The first dataset contained samples collected by ALGO* groups (ALGO dataset) of the Polar Ecology course to describe activities in Svalbard (Petuniabukta) as example of data from the polar regions. The second dataset was collected in Bulgaria (BG dataset) as example of data from the temperate regions. The Bulgarian samples were collected in streams in the vicinity of Temnata Dupka cave (43° 05' 18" N, 23° 22' 59.6" E) and in the shallow ponds with water plants in the Botanical Garden of the Bulgarian Academy of Science in Sofia (42° 38' 41.8" N, 23° 18' 04.4" E). Each dataset was characterized by number of samples collected, number of sites visited, number of taxa observed. Each dataset was characterized by percentages of habitats and communities sampled and taxa at class level observed.

For multivariate analyses, the habitats were grouped into environments according to Elster (2002) with slight modifications (Table 1).

The cyanobacteria and algae were determined according to Ettl et Gärtner (2014), Komárek et Anagnostidis (1999, 2005), Komárek (2013) and Sládeček et Sládečková (1996).

Environment	Habitat
Cryo-environment	Glacier
	Snow/ice
Lacustrine environment	Lake
Lotic environment	River
	Stream
Hydro-terrestrial environment	Intertidal zone
	Seepage
	Shallow pools
	Waterfall
	Wet hummock/thufur tundra
	Wetland
Terrestrial environment	Aerophytic
	Soil
	Stone

Table 1. Grouping of habitats into individual environments. Modified from Elster (2002).

Statistics

The multivariate analysis was performed using a CANOCO 5 software (Ter Braak et Šmilauer 2012). The unconstrained canonical analysis (DCA) and constrained canonical analysis (CCA) were used to test the effect of environment types specified in Table 1 on species/taxa composition as test of database capabilities for ecological research; due to low number of samples in BG dataset, these analyses were performed only for ALGO dataset (*n* = 143). Permutation test for all axes was used to evaluate the statistical significance of the effects of explanatory variable. The results were considered significant for $P < 0.05$.

Results

SampleDTB statistics

At present (December 2015), the Sample DTB contains records on total of 362 samples from 147 sites. Total of 245 taxa were observed.

The ALGO dataset (2011 - 2014) contained records on total of 188 samples from 94 sites. Total of 216 taxa were observed. Majority of habitats visited by ALGO* groups (Fig. 1a) belonged to lotic environments (34.4%), namely streams (32.3%). Hydroterrestrial environments were represented in 26.5% of samples. Surprisingly, very low attention was paid to cryo-environments (4.2%). The percentage of sampled habitat types varied among the years depending on weather conditions during summer season, actual weather conditions during field work, security issues and scientific focus of each ALGO group. The most frequently sampled communities were layer-forming communities like periphyton (37.2%), epilithon (14.4%) and biofilms (13.8%). As biofilms were considered thin (less than 0.5 mm) layers of biomass only (Kvíderová 2011) (Fig. 1b).

The most dominant classes observed in ALGO dataset (Fig. 1c) were diatoms (38.1% of observations; Bacillariophyceae, Coscinodicophyceae, Fragilariophyceae) and cyanobacteria (35.4%). The most frequently observed taxa (Table 2) were filamentous cyanobacteria *Phormidium* sp., *Leptolygbya* sp., and *Pseudanabaena* sp., colony-forming heterocytous *Nostoc* sp., diatoms *Cymbella* sp., *Hannaea arcus*, *Navicula* sp., *Pinnularia* sp. and *Meridion circulare*, and desmid *Cosmarium* sp. (Figs. 2 and 3).

The DCA with environment types as supplementary variable revealed total variation of 8.92148. The supplementary variable explained 18.4%. The adjusted explained variation was 9.0% (Fig. 4, Table 3). The CCA with environment types (pseudo-F=2.5, P=0.002 for all axes) showed total variation of 8.92148. The explanatory variable accounted for 7.7% of total variation. Adjusted explained variation was 4.7% (Fig. 5, Table 5). When the samples were plotted against the environment type, the lotic and cryo- environments separated from terrestrial, hydro-terrestrial and lacustrine ones (Fig. 6).

The BG dataset (2011-2014) contained records on total of 16 samples from 4 sites. Total of 40 taxa were observed. Majority of habitats sampled included streams (43.5 %) and shallow pools (37.5%). The communities sampled (Fig. 4b) were restricted to periphyton (75.0%), epiphyton (18.8%) and epilithon (6.2%). The lower diversity of habitats and communities reflected limited number of sampling sites. Similarly to ALGO dataset, the most dominant classes in BG dataset (Fig. 4c) were Cyanobacteria (37.5%) and diatoms (26.9%). The most frequently observed taxa (Table 2) were cyanobacteria *Phormidium* sp. and *Aphanothece* sp., and diatoms *Melosira* sp., *Cocconeis* sp. and *Navicula* sp. Other species were observed only once, reflecting thus low number of samples (Figs. 5 and 6).

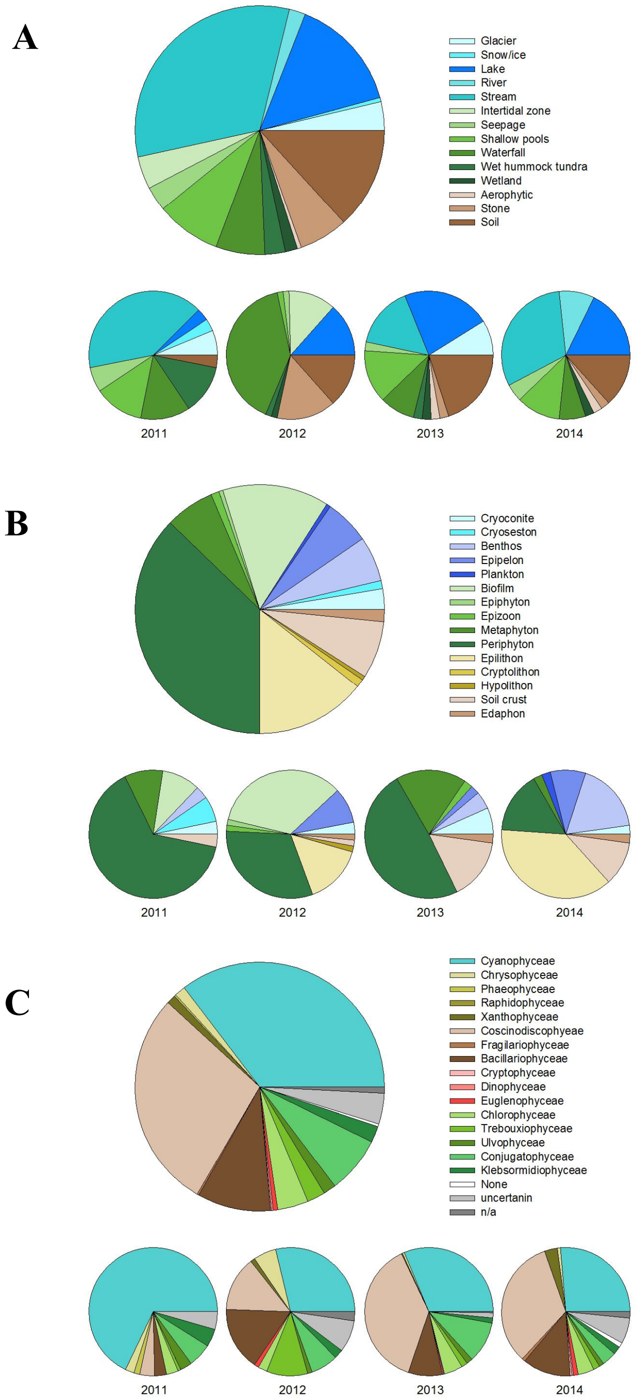

Fig. 1. ALGO dataset. (a) The habitats sampled (b) communities sampled, (c) classes observed. None – no taxon observed, uncertain – not identified, n/a – other microorganisms than algae and cyanobacteria.

	ALGO dataset					BG dataset
	ALGO all	ALGO 2011	ALGO 2012	ALGO 2013	ALGO2014	
1	*Phormidium* sp. (38)	*Phormidium* sp. (10)	*Phormidium* sp. (9)	*Phormidium* sp. (19)	*Hannaea arcus* (12)	*Cocconeis* sp. (4)
2	*Cymbella* sp. (31)	*Leptolyngbya* sp. (8)	*Hannaea arcus* (7)	*Cymbella* sp. (18)	*Microcoleus* sp. (10)	*Phormidium* sp. (3)
3	*Hannaea arcus* (27)	*Schizothrix facilis* (8)	*Prasiola* sp. (7)	*Denticula* sp. (15)	*Cymbella* sp. (10)	*Apanothece* sp. (2)
4	*Leptolyngbya* sp. (26)	*Klebsormidium* sp. (6)	*Meridion* sp. (6)	*Nostoc* sp. (14)	*Meridion circulare* (8)	*Melosira* sp. (2)
5	*Navicula* sp. (23)	*Nostoc commune* (5)	*Navicula* sp. (6)	*Pinnularia* sp. (14)	*Leptolyngbya* sp. (8)	*Navicula* sp. (2)
6	*Pinnularia* sp. (21)	*Phormidesmis* sp. (4)	*Hydrurus foetidus* (5)	*Zygnema* sp. (11)	*Pseudanabaena* sp. (8)	*
7	*Meridion circulare* (19)	*Pseudanabaena* sp. (4)	*Leptolyngbya* sp. (4)	*Meridion circulare* (11)	*Pinnularia* sp. (6)	
8	*Pseudanabaena* sp. (19)	*Hydrurus foetidus* (3)	*Klebsormidium* sp. (2)	*Cosmarium* sp. (11)	*Navicula* sp. (6)	
9	*Nostoc* sp. (18)	*Meridion* sp. (3)	*Mesotaenium* sp. (2)	*Navicula* sp. (10)	*Diatoma* sp. (6)	
10	*Cosmarium* sp. (17)	*Chlamydomonas nivalis* (3)	*Cosmarium* sp. (2)	*Gloeocapsa* sp. (10)	*Chlamydomonas* sp. (5)	

*Other taxons were observed only in one sample.

Table 2. Top ten of mostly observed taxa. Number in parentheses indicates number of observations.

	Axis 1	**Axis 2**	**Axis 3**	**Axis 4**
Eigenvalues	0.7828	0.5555	0.3905	0.3169
Explained variation (cumulative)	8.77	15.00	19.38	22.93
Gradient length	5.58	6.29	5.88	5.97
Pseudo-canonical correlation (suppl.)	0.6385	0.6338	0.4711	0.5218
Case scores centroid	1.56922	3.85645	3.20178	3.23399
Extracted from suppl. data	0.04340	0.05131	0.02711	0.03690

Table 3. The summary of DCA with environment type supplementary variable.

	Axis 1	**Axis 2**	**Axis 3**	**Axis 4**
Eigenvalues	0.3041	0.2445	0.0838	0.0542
Explained variation (cumulative)	3.41	6.15	7.09	7.70
Pseudo-canonical correlation (suppl.)	0.7332	0.7023	0.5030	0.4399
Explained fitted variation (cumulative)	44.30	79.91	92.11	100.00
Extracted from expl. data	0.26600	0.21970	0.24900	0.26529

Table 4. The results of CCA with environment type as explanatory variable.

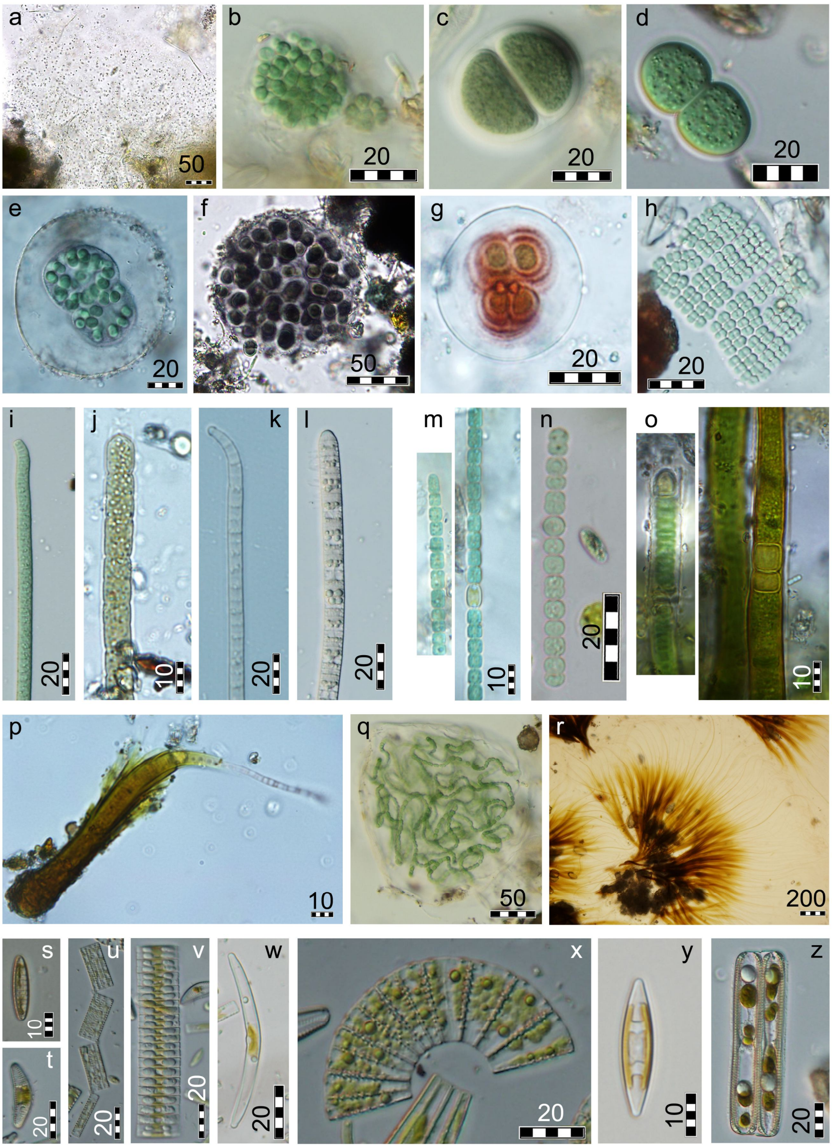

Fig. 2. The representatives of cyanobacteria and diatoms observed by ALGO* groups. The number at the scale indicates scale length in µm.
Coccal types: (a) *Anathece* sp., (b) *Aphanothece* sp., (c) *Chroococcus tenuis*, (d) *Cyanothece* sp., (e) *Gloeocapsa alpina*, (f) *Gloeocapsa athrata*, (g) *Gloeocapsa sanguinea*, (h) *Merismopedia* sp.
Oscillatorian types: (i) *Kamptonema* sp., (j) *Komvophoron* sp., (k) *Microcoleus* sp. (l) *Tychonema* cf. *bornetii*
Heterocytous types: (m) *Anabaena* cf. *jonsonii*, (n) *Anabaena* sp., (o) *Scytonematopsis starmachii*, (p) *Calotrhix* sp., (q) *Nostoc* sp., (q) *Rivularia* sp.
Diatoms: (s) *Denticula* sp. (Bacillariophyceae), (t) *Encyonema* sp. (Bacillariophyceae), (u) *Diatoma tenue* (Fragilariophyceae), (v) *Fragilaria* sp. (Fragilariophyceae), (w) *Hannaea arcus* (Bacillariophyceae), (x) *Meridion circulare* (Fragilariophyceae), (y) *Navicula* sp. (Bacillariophyceae), (z) *Pinnularia* sp. (Bacillariophyceae).

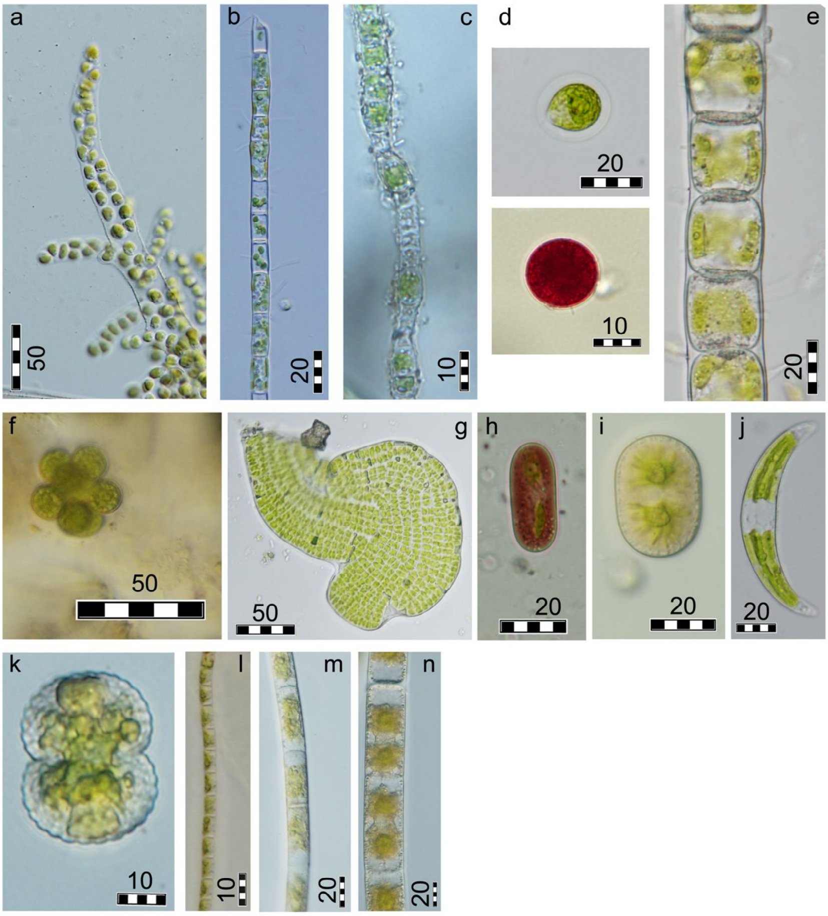

Fig. 3. The representatives of golden, yellow-green and green algae *sensu lato* observed by ALGO* groups. The number at the scale indicates scale length in µm.
Chrysophyceae: (a) *Hydrurus foetidus*
Xanthophyceae: (b) *Tribonema* cf. *vulgare*
Chlorophyta: (c) *Binuclearia* sp. (Ulvophyceae), (d) *Chlamydomonas* cf. *nivalis* (Chlorophyceae), green flagellate and red cyst, (e) *Ulothrix* sp. (Ulvophyceae), (f) *Coelastrum* sp. (Chlorophyceae), (g) *Prasiola* sp. (Trebouxiophyceae)
Streptophyta: (h) *Ancyonema nordenskioeldii* (Conjugatophyceae) (i) *Cylindrocystis brebissonii* (Conjugatophyceae), (j) *Closterium* sp. (Conjugatophyceae), (k) *Cosmarium* sp. (Conjugatophyceae), (l) *Klebsormidium* sp. (Klebsormidiuphyceae), (m) *Mougeotia* sp. (Conjugatophyceae), (n) *Zygnema* sp. (Conjugatophyceae).

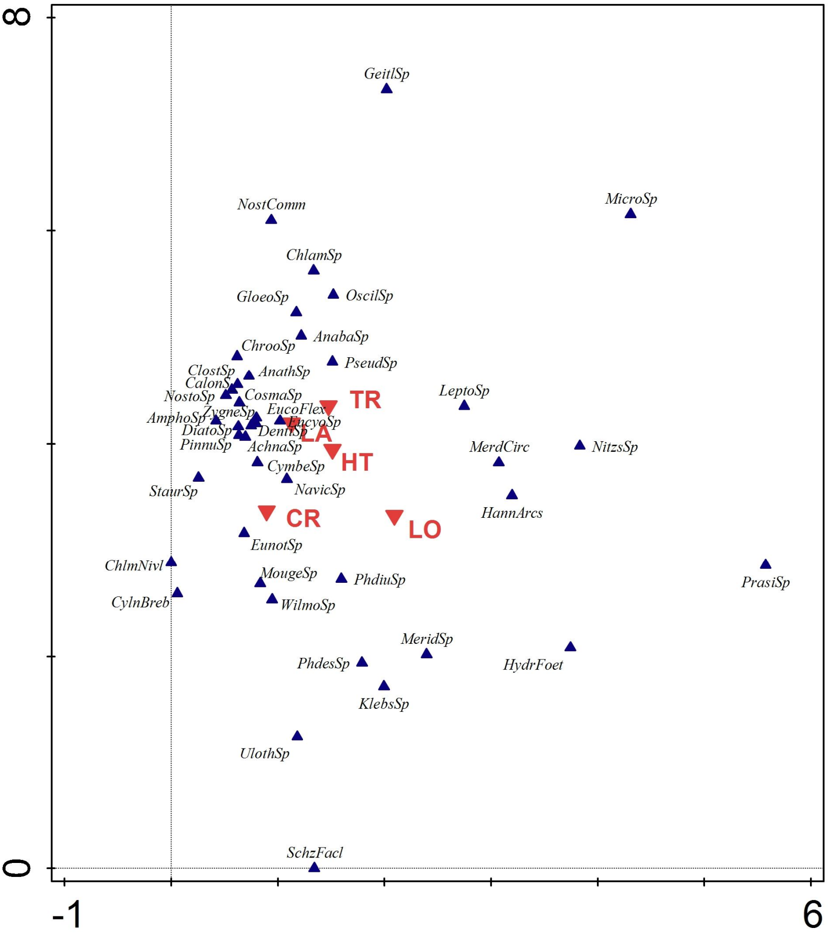

Fig. 4. The DCA diagram of species with environment types as supplementary variable on the first and second ordination axes.

Environment abbreviatons: CR – cryoenvironment, HT- hydroterrestial environment, LA – lacustrine environment, LO – lotic environment, TR – terrestrial environment.

Species abbreviation: *AchnaSp* – *Achnanthidium* sp., *AmphoSp* – *Amphora* sp., *AnabaSp* – *Anabaena* sp., *AnathSp* – *Anathece* sp., *CalonSp* – *Caloneis* sp., *ChlamSp* – *Chlamydomonas* sp., *ChlmNivl* – *Chlamydomonas* cf. *nivalis*, *ChrooSp* – *Chroococcus* sp., *ClostSp* – *Closterium* sp., *CosmaSp* – *Cosmarium* sp., *CylnBreb* – *Cylindrocystis brebissonii*, *CymbeSp* – *Cymbella* sp., *DentiSp* – *Denticula* sp., *DiatoSp* – *Diatoma* sp., *EncyoSp* – *Encyonema* sp., *EucoFlex* – *Eucocconeis flexella*, *EunotSp* – *Eunotia* sp., *GeitlSp* – *Geitletinema* sp., *GloeoSp* – *Gloeocapsa* sp., *HannArcs* – *Hannaea arcus*, *HydrFoet* – *Hydrurus foetidus*, *KlebsSp* - *Klebsormidium* sp., *LeptoSp* – *Leptolyngbya* sp., *MerdCirc* – *Meridion circulare*, *MeridSp* – *Meridion* sp., *MicroSp* – *Microcoleus* sp., *MougeSp* – *Mougeotia* sp., *NavicSp* – *Navicula* sp., *NitzsSp* – *Nitzschia* sp., *NostComm* – *Nostoc commune*, *NostoSp* – *Nostoc* sp., *OscilSp* – *Oscillatoria* sp., *PhdesSp* – *Phormidesmis* sp., *PhdiuSp* – *Phormidium* sp., *PinnuSp* – *Pinnularia* sp., *PrasiSp* – *Prasiola* sp., *PseudSp* – *Pseudanabaena* sp., *SchzFacl* – *Schizothrix facilis*, *StaurSp* – *Stauroneis* sp., *UlothSp* – *Ulothrix* sp., *WilmoSp* – *Wilmottia* sp., *ZygneSp* – *Zygnema* sp.

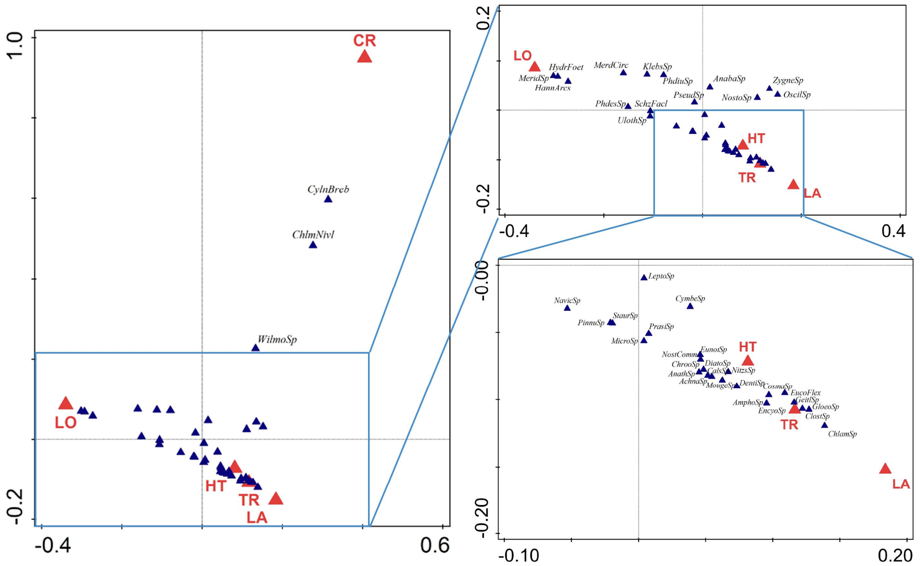

Fig. 5. The CCA diagram of species with environment types as explanatory variable on the first and second ordination axes.
Environment abbreviatons: CR – cryoenvironment, HT- hydroterrestial environment,
LA – lacustrine environment, LO – lotic environment, TR – terrestrial environment.
Species abbreviation: *AchnaSp* – *Achnanthidium* sp., *AmphoSp* – *Amphora* sp., *AnabaSp* – *Anabaena* sp., *AnathSp* – *Anathece* sp., *CalsSp* – *Caloneis* sp., *ChlamSp* – *Chlamydomonas* sp., *ChlmNivl* – *Chlamydomonas* cf. *nivalis*, *ChrooSp* – *Chroococcus* sp., *ClostSp* – *Closterium* sp., *CosmaSp* – *Cosmarium* sp., *CylnBreb* – *Cylindrocystis brebissonii*, *CymbeSp* – *Cymbella* sp., *DentiSp* – *Denticula* sp., *DiatoSp* – *Diatoma* sp., *EncyoSp* – *Encyonema* sp., *EucoFlex* – *Eucocconeis flexella*, *EunotSp* – *Eunotia* sp., *GeitlSp* – *Geitletinema* sp., *GloeoSp* – *Gloeocapsa* sp., *HannArcs* – *Hannaea arcus*, *HydrFoet* – *Hydrurus foetidus*, *KlebsSp* - *Klebsormidium* sp., *LeptoSp* – *Leptolyngbya* sp., *MerdCirc* – *Meridion circulare*, *MeridSp* – *Meridion* sp., *MicroSp* – *Microcoleus* sp., *MougeSp* – *Mougeotia* sp., *NavicSp* – *Navicula* sp., *NitzsSp* – *Nitzschia* sp., *NostComm* – *Nostoc commune*, *NostoSp* – *Nostoc* sp., *OscilSp* – *Oscillatoria* sp., *PhdesSp* – *Phormidesmis* sp., *PhdiuSp* – *Phormidium* sp., *PinnuSp* – *Pinnularia* sp., *PrasiSp* – *Prasiola* sp., *PseudSp* – *Pseudanabaena* sp., *SchzFacl* – *Schizothrix facilis*, *StaurSp* – *Stauroneis* sp., *UlothSp* – *Ulothrix* sp., *WilmoSp* – *Wilmottia* sp., *ZygneSp* – *Zygnema* sp.

Discussion

The ALGO dataset provided interesting insight in inter-annual variability in habitats and communities sampled and taxonomical classes observed. The types of habitats and communities sampled were probably affected by (a) actual and long-term weather conditions in case of temporary snow fields inhabited mainly by *Chlamydomonas* cf. *nivalis* (Kvíderová 2012) that fields were not available every year due to an early snow-melt and/or late arrival of ALGO group; (b) security issues, for instance polar bears in vicinity of sampling sites or actual weather conditions out of limits for safe transport on zodiac; and (c) scientific focus of each ALGO group, *e.g.* preference of hydro-terrestrial habitats in 2012 or cyanobacteria in 2011.

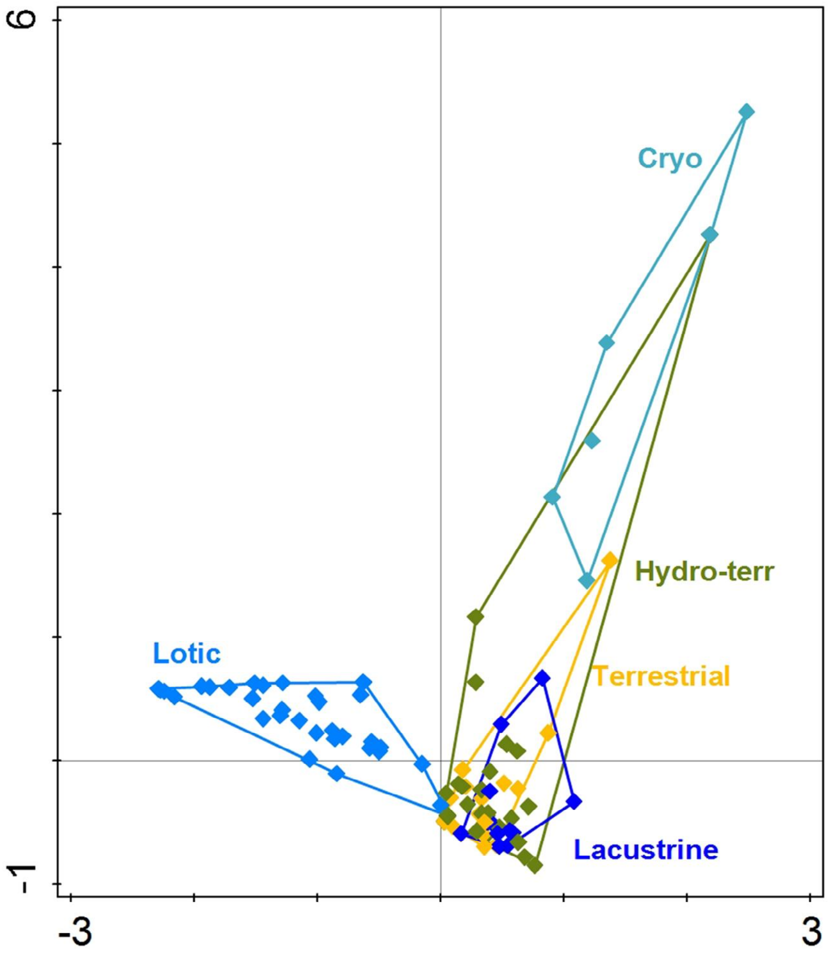

Fig. 6. The sample groups recognized by CCA analysis on the first and second ordination axes.

Cyanobacteria, diatoms and green algae *sensu lato* (*i.e.* Chlorophyta and Streptophyta) were dominant photosynthetic microorganisms that is in agreement with other ecological studies from Svalbard (Kim et al. 2008, Komárek et al. 2012, Lesniak 2012, Matula et al. 2007, Pinseel et al. 2014, Skulberg 1996). Some sites were sampled several times, so their taxonomical diversity could be compared. In some sites, the community structure did not change during four consecutive years. For instance, the microbial community in lower reach of a stream with waterfall in a steep valley on the eastern bank of Petuniabukta was always dominated by golden alga *Hydrurus foetidus* and diatoms *Hannaea arcus* and *Meridion circulare*. Similarly green alga *Prasiola* sp. was always found growing on nutrient rich soils under bird cliffs. On the other hand, interseasonal changes were apparent as well. Green filamentous algae *Zygnema* sp. and *Klebsormidium* sp. dominated in shallow streams between 2011 and 2013, and they were replaced by green-yellow filaments of *Tribonema* ssp. in 2014.

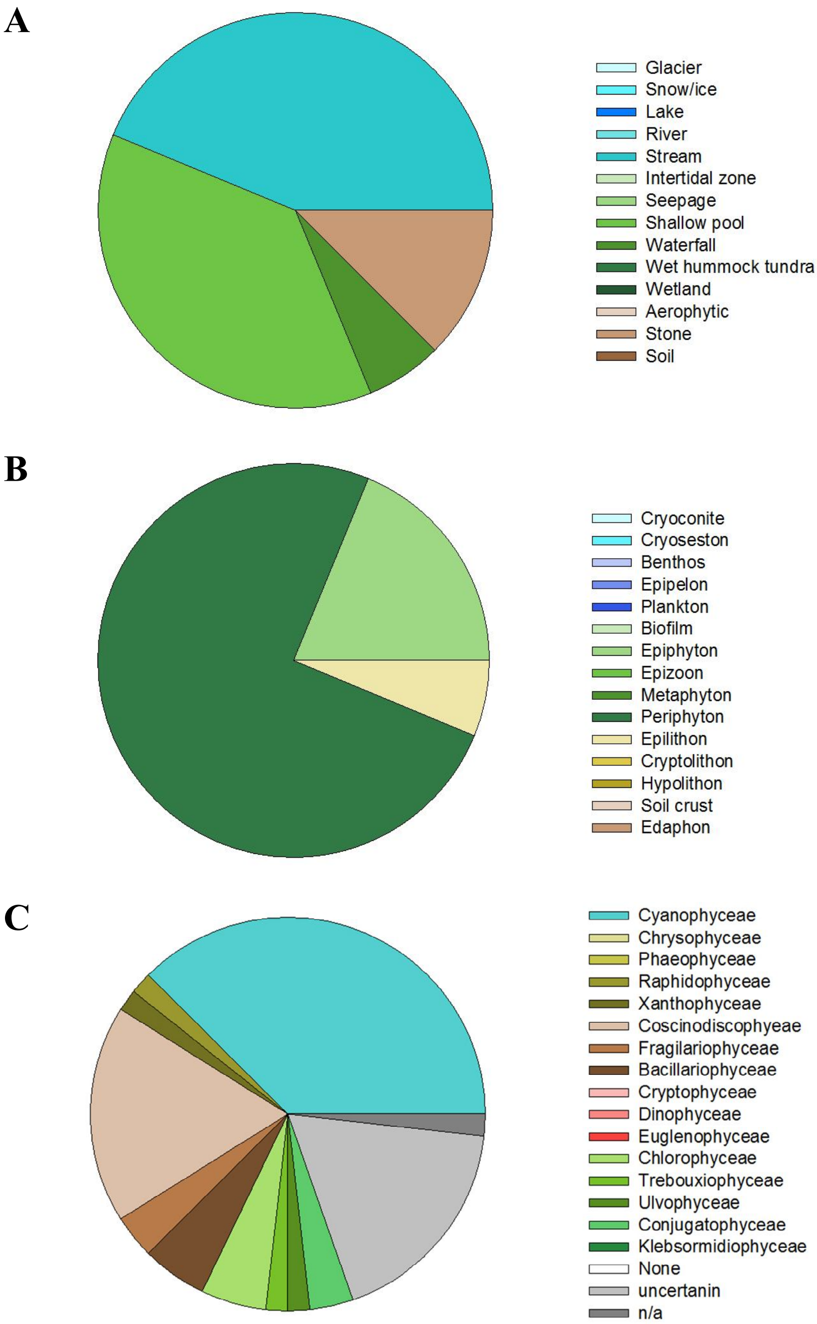

Fig. 7. BG dataset. (a) The habitats sampled (b) communities sampled, (c) classes observed. None – no taxon observed, uncertain – not identified, n/a – other microorganisms than algae and cyanobacteria.

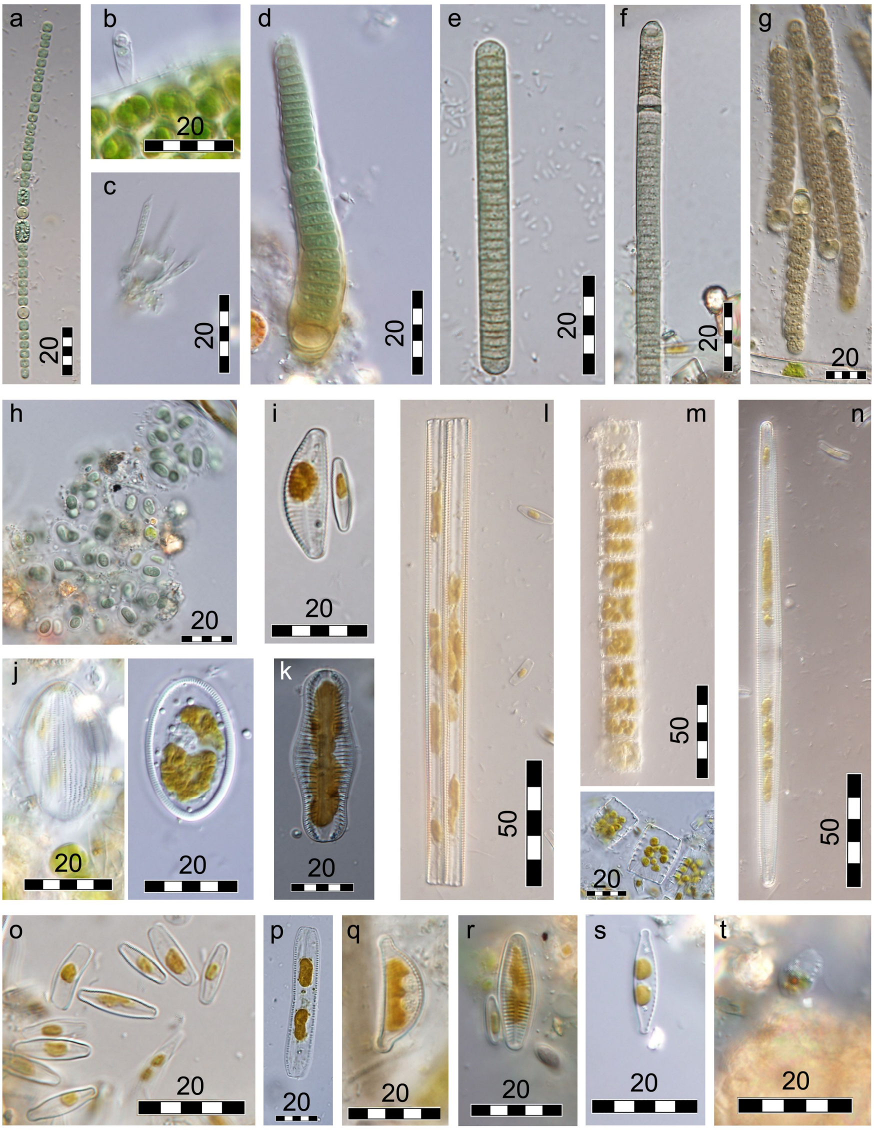

Fig. 8. The representatives of cyanobacteria, diatoms, and a raphidophycean alga observed in BG dataset. The number at the scale indicates scale length in µm.
Cyanophyceae: (a) *Anabaena* cf. *oscillarioides*, (b) *Chamaesiphon* cf. *minutus*, (c) *Clastidium* sp., (d) *Dichothrix* sp., (e) *Oscillatoria* cf. *rupicola*, (f) *Phormidium* sp., (g) *Rivularia aquatica*, (h) *Gloeothece* cf. *membranacea*
Diatoms: (i) *Cymbella* sp. (Bacillariophyceae), (j) *Cocconeis* cf. *placentula* (Bacillariophyceae), (k) *Fragilaria* sp. (Fragilariophyceae), (l) *Diatoma* sp., (Fragilariophyceae) (m) cf. *Melosira* sp. (Coscinodiscophyceae), (n) cf. *Synedra* sp. (Fragilariophyceae), (o) cf. *Achnanthidium* sp. (Bacillariophyceae), (p) cf. *Epithemia adnata* (Bacillariophyceae), (q) *Encyonema* sp. (Bacillariophyceae), (r) *Navicula* sp. (Bacillariophyceae), (s) *Nitzschia* cf. *dissipata* (Bacillariophyceae)
Raphidophyceae: (t) *Gonyostomium* sp.

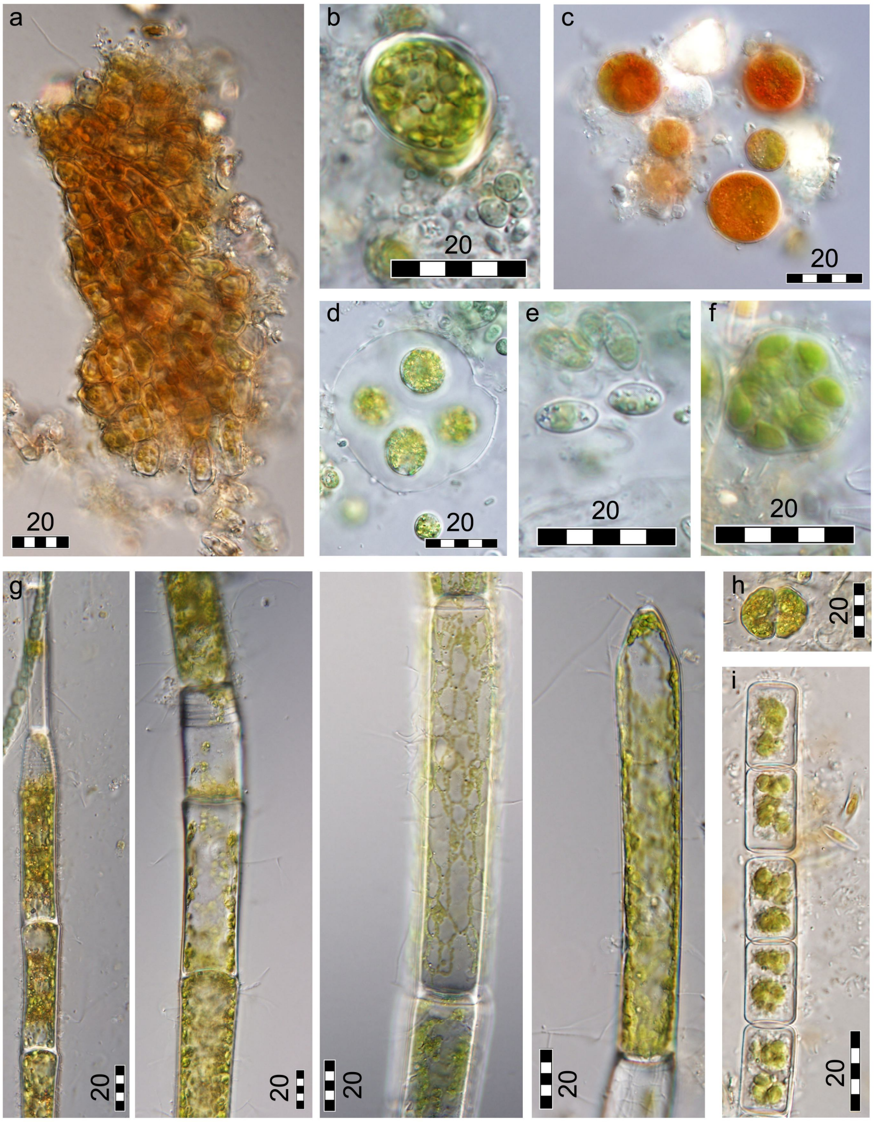

Fig. 9. The representatives of yellow-green and green algae *sensu lato* observed by ALGO* groups. The number at the scale indicates scale length in µm.
Xanthophyceae: (a) cf. *Heteropedia polychloris*
Chlorophyta: (b) *Bracteacoccus* sp. (Chlorophyceae), (c) *Haematococcus pluvialis* (Chlorophyceae), (d) cf. *Asterococcus* sp. (Chlorophyceae), (e) *Oocystis* sp. (Trebouxiophyceae), (f) *Tetracystis* sp. (Chlorophyceae), (g) *Oedogonium* sp. (Chlorophyceae)
Streptophyta: (h) *Cosmarium* sp. (Conjugatophyceae), (i) *Mougeotia* sp. (Conjugatophyceae)

According to their taxonomical composition, the lotic environments and partly cryo-environments were separated from terrestrial, hydro-terrestrial and lacustrine ones (Fig. 6). Overlapping of terrestrial, hydro-terrestrial and lacustrine ones was caused probably by selection of sampling point. Since there are continual spatiotemporal gradients of water availability (Elster 2002) in majority of habitats, the actual microhabitat sampled may correspond to other habitat/community. For instance, if a sample is taken from partly flooded lake littoral with dominant mosses, the genera and species observed may be similar to surrounding wet tundra rather than to benthic communities of littoral with stony bottom. Additional physico-chemical data (pH, temperature, *etc.*) could further separate these environment, however, these data are incomplete and do not allow statistical analyses, recently.

Despite of its small size, the BG dataset shows diverse communities of cyanobacteria and algae.

At present (December 2015), the BG dataset is not suitable for multivariate analyses due to low number of samples and very limited habitat types and communities sampled. Especially, uniqueness of each sample in term of species composition restricts such analyses. In future, we plan to extend the BG dataset by sampling in various extreme environments in Bulgaria. We would like to focus on acidic environments, thermal and mineral springs and alpine snow habitats.

Dataset comparisons

The ALGO dataset was much larger than BG dataset in total numbers of sample collected and sites visited. However, when these data are normalized to person-year, the efficiency of sample collecting is higher in BG dataset, since there are 5.69 samples per person and a year in ALGO dataset, and 8 samples per person and a year in BG dataset. The number of sites visited is slightly higher in the ALGO dataset, 2.84 sites per person and a year in ALGO compared to 2 in BG dataset. However, the number of samples per site was higher in the BG dataset than in the ALGO dataset indicating higher diversity of habitats/communities per site in Bulgaria.

Conclusions

In the first operational test, SampleDTB proved functional. The database provided all requested outputs. The functional test revealed some incomplete sample data, like missing zone data in some BG samples which could be easily filled in. However, some missing data, like physico-chemical analyses of the sites of collection, are not available, and therefore limit the type of statistical analyses as discussed in Kvíderová (2014). Low number of sample in selected dataset, like in the BG one, also restricts data utilization in statistical analyses. Despite of these limitations, the ALGO dataset demonstrated that ecological studies based on SampleDTB data are possible and such approach might be helpful for field algologists.

Moreover, the database may be useful for isolation of novel algal and cyanobacterial strains, and and in ecophysiological studies focused on adaptation/acclimatization strategies providing detailed knowledge of the conditions at the original locality. Data on physico-chemical parameters of original locality can contribute to selection of appropriate cultivation technique and cultivation conditions. For instance, a soil alga should be cultivated preferably on solid medium and in low irradiances. Knowledge of sampling dates may contribute to calculation of time spent by a selected algal in culture collection conditions, and hence approximate number of generations, in order to estimate the effects of acclimation *sensu* Elster (1999) on strain ecophysiological requirements. The longer time spent in culture collection condition, the higher probability of mutation occurs and selection of genotypes better adapted to cultivation conditions could lead to shift in strain growth requirements. Data on original *in situ* condition may contribute to explanation of results of laboratory experiments. For example, optimum growth temperature of Svalbard strain *Stichococcus* Kováčik 1988/9 was higher than optimum growth temperature of other polar strains used in the study, since the strain *Stichococcus* Kováčik 1988/9 originated from warm Troll Springs while the other strains from periglacial soil (Kvíderová et Lukavský 2005). Data on a species/genera observation(s) in the field may reveal its tolerance range and thus may contribute to taxonomical identification (Komárek 2010).

References

ELSTER, J. (1999): Algal versatility in various extreme environments. *In*: J. Seckbach (eds.): Enigmatic microorganisms and life in extreme environments. Kluwer Academic Publishers, pp. 215-227.

ELSTER, J. (2002): Ecological classification of terrestrial algal communities in polar environments. *In*: L. Beyer, M. Bötler (eds.): Geoecology of Antarctic ice-free coastal lanscapes. Springer-Verlag, pp. 303-326.

ETTL, H., GÄRTNER, G. (2014): Syllabus der Boden-, Luft- und Flechtenalgen. Springer, Berlin Heidelberg, 773 p.

JONES, M. B., SCHILDHAUER, M. P., REICHMAN, O. J. and BOWERS, S. (2006): The new bioinformatics: integrating ecological data from the gene to the biosphere. *Annual Review of Ecology, Evolution, and Systematics,* 37: 519-544.

KHALSA, S. J. S., YARMEY, L. (2014): Data management challenges in polar ecology. *In*: J. Kavan, A. Bernardová (eds.): Polar Ecology Conference 2014. Faculty of Science, University of South Bohemia in České Budějovice, České Budějovice, 70 p.

KIM, G. H., KLOCHKOVA, T. A. and KANG, S. H. (2008): Notes on freshwater and terrestrial algae from Ny-Ålesund, Svalbard (high Arctic sea area). *Journal of Environmental Biology,* 29: 485-491.

KOMÁREK, J., ANAGNOSTIDIS, K. (1999): Süßwasserflora von Mitteleuropa 19/1. Cyanoprokaryota. 1.Teil: Chroococcales. Gustav Fischer Verlag, Jena, 548 p.

KOMÁREK, J., ANAGNOSTIDIS, K. (2005): Süßwasserflora von Mitteleuropa 19/2. Cyanoprokaryota. 2.Teil: Oscillatoriales. Elsevier/Spektrum, Heidelberg, 759 p.

KOMÁREK, J. (2010): Recent changes (2008) in cyanobacteria taxonomy based on a combination of molecular background with phenotype and ecological consequences (genus and species concept). *Hydrobiologia,* 639: 245-259.

KOMÁREK, J., KOVÁČIK, L., ELSTER, J. AND KOMÁREK, O. (2012): Cyanobacterial diversity of Petuniabukta, Billefjorden, central Spitzbergen. *Polish Polar Research,* 33: 347-368.

KOMÁREK, J. (2013): Süßwasserflora von Mitteleuropa 19/3. Cyanoprokaryota. 3.Teil: Heterocytous genera. Springer, Heildelberg, 1131 p.

KVÍDEROVÁ, J., LUKAVSKÝ, J. (2005): The comparison of ecological characteristics of *Stichococcus* (Chlorophyta) strains isolated from polar and temperate regions. *Algological Studies,* 118: 127-140.

KVÍDEROVÁ, J. (2011): Biofilm. *In*: M. Gargaud, R. Amils, J. C. Quintanilla, H. J. Cleaves II, W. M. Irwine, D. L. Pinti and M. Viso (eds.): Encyclopedia of astrobiology. Springer, Berlin Heildelberg, pp. 170-171.

KVÍDEROVÁ, J. (2012): Research on cryosestic communities in Svalbard: the snow algae of temporary snowfields in Petuniabukta, Central Svalbard. *Czech Polar Reports,* 2: 8-19.

KVÍDEROVÁ, J., ELSTER, J. (2013): Standardized algal growth potential and/or algal primary production rates of maritime Antarctic stream waters (King George Island, South Shetlands). *Polar Research,* 32: 17.

KVÍDEROVÁ, J. (2014): Sample database of the Centre for Polar Ecology - Database design and data management. *Czech Polar Reports,* 4: 140-148.

LESNIAK, V. (2012): Diversity and ecophysiological performance of cyanobacteria in wet meadow, Petunia Bay, Central Svalbard, Universite Blaise Pascal, Clermont-Ferrand, 39 p.

MATULA, J., PIETRYKA, M., RICHTER, D. and WOJTUN, B. (2007): Cyanoprokaryota and algae of Arctic terrestrial ecosystems in the Horsund area, Spitzbergen. *Polish Polar Research,* 28: 283-315.

PINSEEL, E., KOPALOVÁ, K., VAN DEN VIJVER, B. and LANGE-BERTALOT, H. (2014): Freshwater diatoms of Petunia Bay (Spitzbergen). Universiteit Antwerpen, Botanical Garden Meise, University of South Bohemia, Antwerpen, Meise, České Budějovice, 98 p.

SKULBERG, O. M. (1996): Terrestrial and limnic algae and cyanobacteria. *In*: A. Elvebakk, P. Prestrud (eds.): A Catalogue of Svalbard Plants, Fungi, Algae and Cyanobacteria. Norsk Polarinstitutt Skrifter, pp. 383-395.

SLÁDEČEK, V., SLÁDEČKOVÁ, A. (1996): Atlas vodních organismů se zřetelem na vodárenství, povrchové vody a čistírny odpadních vod. 1. díl: Destruenti a producenti. [Atlas of aquatic organisms with respect to water supply, surface waters and wastewater treatment plants. Part 1. Decomposers and producers]. Agrospoj, Praha, 350 p.

TER BRAAK, C. J. F., ŠMILAUER, P. (2012): Canoco reference manual and user's guide: software for ordination, version 5.0. Microcomputer Power, Ithaca, USA, 496 p.

Web sources / Other sources

Report on Polar Ecology course Svalbard 2011. Centre for Polar Ecology, Faculty of Science, University of South Bohemia, České Budějovice, 20 p.
http://polar.prf.jcu.cz/data/polar%20ecology.pdf

Report on Polar Ecology course Svalbard 2012. Centre for Polar Ecology, Faculty of Science, University of South Bohemia, České Budějovice, 28 p.
http://polar.prf.jcu.cz/data/polar_ecology_course_report_2012.pdf

Report on Polar Ecology course Svalbard 2013. Centre for Polar Ecology, Faculty of Science, University of South Bohemia, České Budějovice, 31 p.
http://polar.prf.jcu.cz/data/Polar_ecology_course_2013.pdf

Report on Polar Ecology course 2014. Centre for Polar Ecology, Faculty of Science, University of South Bohemia, České Budějovice, 26p.
http://polar.prf.jcu.cz/data/PolarEKO2014/ReportTeaching2014final.pdf